

# Criteria for Analysis and Comparison of Experimental Data under Conditions of Uncertainty

Sergey I. Kumkov<sup>1, 2, a)</sup>, Zinaida V. Kataeva<sup>3, b)</sup>, Oleg V. Shilovskikh<sup>3</sup>

<sup>1</sup>*N.N. Krasovskii Institute of Mathematics and Mechanics, Ural Branch of RAS, Yekaterinburg, Russia*

<sup>2</sup>*Yeltsin Ural Federal University, Yekaterinburg, Russia*

<sup>3</sup>*IRTC Eye Microsurgery Ekaterinburg Center, Ekaterinburg, Russia*

<sup>a)</sup>Corresponding author: kumkov@imm.uran.ru

<sup>b)</sup>kataeva@eyeclinic.ru

**Abstract.** The paper deals with investigation of the important problem of processing the ophthalmic data on the post-operation status of patients. The groups of patients differ by the type (technology) of fixing the intraocular lenses (IOL). Validity of each type of technology is estimated by computation of criteria for distinction of data between groups. The initial information comprises measurements of several ophthalmic indices. The samples on each index are very short; in each index, as a rule, the samples of patients' groups overlap each other; any probabilistic characteristics of the measuring indices are unknown; any probabilistic characteristics of the measuring errors are also unknown. So, the standard methods of mathematical statistics can be applied only in the formal way and have shown to be inefficient. In contrast, the Hausdorff distance (from the set theory) as the criterion of distinction between two samples (both for one- and, especially, for two-dimensional indices) demonstrated to be reliable to distinct the patient's status. Computations of the Hausdorff distance are valid for any relative location of point sets under comparison.

**Keywords:** Experimental ophthalmic data, post operation procession, patients, groups, samples, criteria, estimation, comparison, distinction.

## PROBLEM FORMULATION

In the paper, we describe investigation of application of possible criteria to analysis and decision making in procession of ophthalmic data on the post-operation status between two groups of patients. The groups of patients differ by the type (technology) of fixing the intraocular lenses (IOL). Validity of each type of ophthalmic technology was estimated by computation of certain criterion.

Procession of experimental information is implemented under conditions of strong uncertainty: samples of measurements on each ophthalmic index are very short; on each index, the samples of groups often overlap each other; any probabilistic characteristics of the measuring errors are unknown.

For each group of patients, the experimental data are represented as samples of measurements on each index (type of ophthalmic characteristic)  $Cr(k, m, j)$ ,  $k = 1, K$ ,  $m = 1, M$ ,  $j = 1, J$ , where  $K$  is the number of measured indices;  $M$  is the number of patients in the group;  $J$  is the number of groups;  $Cr(k, m, j)$  is the sample of measurements of the  $k$ -index for the  $m$ -th patient of the group  $j$ .

**Problem formulation:** on the basis of the given experimental data, investigate possible types of criteria for reliable quantitative describing the distinction between groups of patients.

## INVESTIGATED CRITERIA FOR ESTIMATING DISTINCTION BETWEEN GROUPS OF PATIENTS. COMPUTATION RESULTS

The following criteria have been probated: formal application of the "Least Squares Means"-criterion (LSQM); the "Intervals-Medians"-criterion; the "Mann-Whitney U"-index; formal application of the Smirnov, Cramer, Student, Fisher, and Rosenbaum criteria; and the "Hausdorff distance"-criterion. As an example, we consider computations of the mentioned criteria on example of experimental data for two groups of patients on two indices "3astigm" and "5astigm" (Fig. 1, measurements are marked by circles and rectangles, correspondingly).

**1. The “Least Squares Mean”-criterion [2,3].** Since short samples (only 7 measurements) for each Group 1 and 2 and absence of probabilistic characteristics of measuring errors, values (mean and standard deviation) of each criterion are calculated only formally as follows:

$$Mean(k,i) = (\sum_{m=1,7} Cr(k,m,i)) / 7, \quad i = 1,2; \quad \sigma(k,i) = (\sum_{m=1,7} (Cr(k,m,i) - Mean(k,i))^2 / 6,$$

where  $k$  is the index number;  $i$  is the number of the patient group.

Results of computation of this criterion take the following values.

**Index “3astigm”:** Group 1,  $Mean(1,1) = 0.086$ ,  $\sigma(1,1) = 0.097$ ; Group 2,  $Mean(1,2) = 0.063$ ,  $\sigma(1,2) = 0.051$ .

**Index “5astigm”:** Group 1,  $Mean(2,1) = 0.247$ ,  $\sigma(2,1) = 0.152$ ; Group 2,  $Mean(2,2) = 0.098$ ,  $\sigma(2,2) = 0.094$ . These results are shown in Fig. 1.

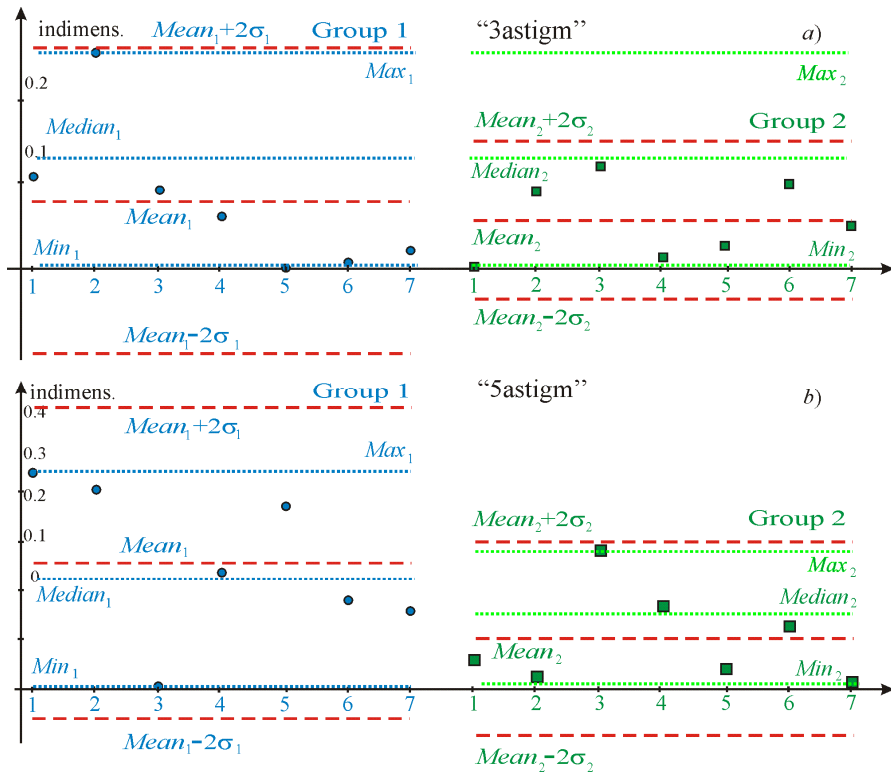
Note absence of visual reliable distinction:  $\pm 2\sigma(1,...)$ -interval absorbs the  $\pm 2\sigma(2,...)$ -interval on both indices in both groups and, moreover, the mean values almost coincide.

**2. The “Intervals--Medians”-criterion [2,3].** In practice, this criterion is one of the simplest approach for comparison of intervals (areas)  $[Min(k, i), Max(k, i)]$  and medians  $Median(k, i)$  of the samples. For the given samples of Groups 1 and 2 this criterion is calculated as follows:

$$k = 1,2, i = 1,2 : Min(k, i) = \min_{m=1,7} \{Cr(k,m, i)\}, \quad Max(k, i) = \max_{m=1,7} \{Cr(k,m, i)\}; \\ Median(k, i) = 0.5 (Min(k, i) + Max(k, i)).$$

Results of computation of this criterion take the following values. **Index “3astigm”:** Group 1,  $Min(1,1) = 0$ ,  $Median(1,1) = 0.14$ ,  $Max(1,1) = 0.28$ ; Group 2,  $Min(1,2) = 0.002$ ,  $Median(1,2) = 0.068$ ,  $Max(1,2) = 0.134$ . **Index “5astigm”:** Group 1,  $Min(2,1) = 0.003$ ,  $Median(2,1) = 0.2125$ ,  $Max(2,1) = 0.422$ ; Group 2,  $Min(2,2) = 0.011$ ,  $Median(2,2) = 0.142$ ,  $Max(2,2) = 0.273$ .

These results are also shown in Fig. 1.



**FIGURE 1.** Samples of measurements and results of computations of the LSQM- and Interval-Medians-criteria. It is seen that  $[Min(...), Max(...)]$ -intervals and medians can not be used to distinct reliably data of Groups 1 and 2.

Moreover, difference in medians on the “3astigm” index is inessential: only 0.072; the similar differences in medians on the “5astigm” index is also inessential: only 0.070. Similarly, comparison of means on the “3astigm” and “5astigm” indices is almost inessential: only 0.149.

So, both the “LSQM”-criterion and the “Intervals–Medians” one demonstrated to be ineffective for quantitative distinguishing the ophthalmic characteristics of patients of Groups 1 and 2.

**3. The “Mann–Whitney U”-index [4,5].** Further, this classical non-parametric index was probated since its wide popularity in processing the medical data. Note that calculation of this criterion is based on introducing special empiric ranging of measurements in the compared samples of measurements.

Computations show that for the whole collection of experimental data [1] (on 14 indices) the typical results on this criterion are **unsatisfactory**. For example, consider Groups 1 and 2 on the index “3astigm”. The critical (maximal admissible value for distinction) level of vicinity for these samples is 8, but the calculation value is 24. So, these two samples are statistically indistinguishable.

The similar result takes place for Group 1 and 2 on the index “5astigm”. The critical level of vicinity for these samples is 8, but the calculation value is 11. So, these two samples are also statistically indistinguishable.

This is the result of the strong overlapping of samples on both indices.

#### **4. The Smirnov, Cramer, Student, Fisher, and Rosenbaum criteria [6–8].**

Special analysis was carried out on possibility of application of these standard classical statistical criteria for estimating the statistical distinction between two samples (Fig. 1). Note that all these methods are also based on hypothesis w.r.t. probability distributions of errors in measured data. So, they can be also applied formally.

But since overlapping and absorption of the span-interval values of the given experimental samples, all these criteria have shown to be absolutely unable to distinct data of the patient Groups 1 and 2 in both indices “3astigm” and “5astigm”.

**5a. The “Hausdorff distance”-criterion for one-dimensional data [9].** For more sophisticated quantitative estimation of distinction between two investigated groups of patients, new ideas from the mathematical theory of sets have been engaged. Namely, it was suggested to characterize distinction between sets of points (*i.e.*, samples of measurements) by the **Hausdorff distance**. Clarify the main computational procedures of this approach. Let two samples of some  $k$ -index be given for Groups 1 and 2:  $Cr(k,m,1)$  and  $Cr(k,n,2) \setminus \}$ ,  $m = 1,7$ ,  $n = 1,7$ . Then the Hausdorff distance  $Hd(k,1-2)$  between these two one-dimensional pointwise sets is calculated as follows:

$$D(k,1-2) = \max_{m=1,7} \{ \min_{n=1,7} \{ |(Cr(k,m,1) - Cr(k,n,2))| \} \},$$

$$D(k,2-1) = \max_{n=1,7} \{ \min_{m=1,7} \{ |(Cr(k,m,1) - Cr(k,n,2))| \} \}, \quad Hd(k,1-2) = D(k,1-2) + D(k,2-1).$$

Here,  $D(k,1-2)$  and  $D(k,2-1)$  can interpreted as estimations of “sizes” of the non-coinciding parts of the sets. This criterion has transparent practical sense: the calculated value  $Hd(\dots)$  is a quantitative estimate for the degree of the “offset–parts” of the sets under comparison.

Note important property of this criterion: its value is equal to 0 for completely coinciding sets. In the case of only one-point-intersection its value grows to value 2. In the case of non-intersection, the value exceeds 2.

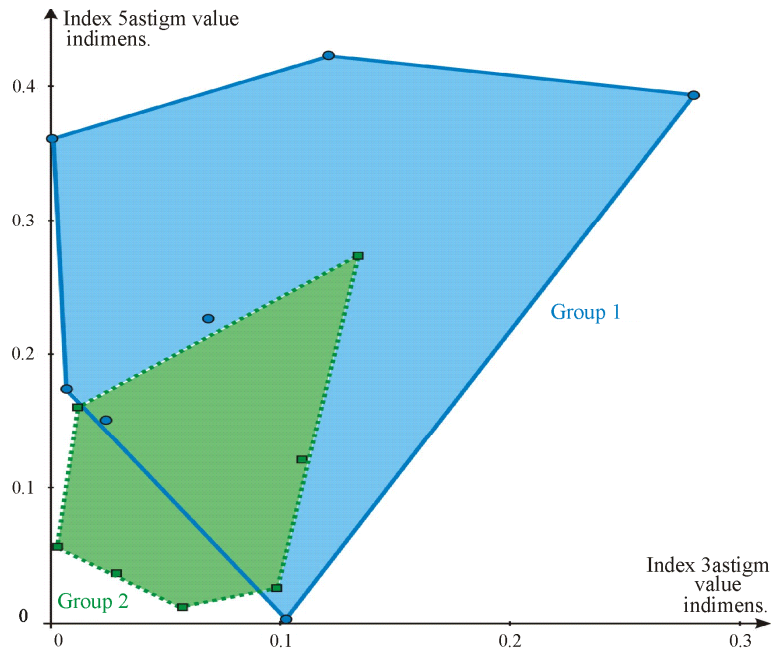
Calculated Hausdorff distance  $Hd(1,1-2)$  between the one-dimensional samples of Groups 1 and 2 on the first index “3atgsm” took the value 0.620. The estimate  $Hd(2,1-2)$  the one-dimensional samples of Groups 1 and 2 on the second index “5atgsm” took the value 0.558.

Ophthalmologists suppose that these values more reliably characterize differences of the ophthalmic properties between corresponding samples of patients in comparison with all earlier mentioned criteria.

**5b. The “Hausdorff distance”-criterion for two-dimensional data.** It was found that more informative estimate of distinction is provided by using the approach to calculation of this criterion been applied to the following two-dimensional vector data. Let us take for Group 1 its two samples on some given indices  $k$  and  $l$ , and take for Group 2 its two samples on the same indices  $k$  and  $l$ :  $\{Cr(k,m,1), Cr(l,m,1), m = 1,7\}$ ;  $\{Cr(k,n,2), Cr(l,n,2)\}$ ,  $n = 1,7$ . Now compare these two-dimensional point sets in the plane  $Cr(k) \times Cr(l)$  by the Hausdorff distance  $Hd(k,l,1-2)$ . Calculation of this criterion is implemented by corresponding formulas [9].

Note important details of the Hausdorff approach in comparison with the mentioned statistical approaches and non-parametric ones: no *any information on probabilistic properties* of the measured data is used; if the both two-dimensional sets comprise the same samples and completely coincide, then the distance (*i.e.*, difference)  $Hd(k,l,1-2)$  is reliably equal zero; if the sets *do not overlap each other*, then the distant  $Hd(k,l,1-2)$  exceeds the value 2.

Picture of two-dimensional samples in the plane of indices 3astm $\times$ 5astm is shown in Fig. 2.



**FIGURE 2.** Two-dimensional sets of samples in the plane of indices 3astm×5astm.

The reliable visual distinction between two groups of patients is seen. Here, the minimal outer envelopes of the two-dimensional sets are shown conditionally. The Hausdorff distance took the value 0.794 that reliably exceeds the values 0.620 and 0.558 for the mentioned cases of one-index analysis.

## CONCLUSION

Analysis of distinction between groups of patients was performed on the basis of the real post-operation ophthalmic data. As a new criterion for comparison data between groups of patients, the Hausdorff distance was probated. This criterion demonstrated to be significantly more effective and reliable for distinction overlapping samples in the cases of one-dimensional, and, especially, of two-dimensional (and multi-dimensional) data.

## ACKNOWLEDGMENTS

The work was supported by the RFBR grant, project no. 18-01-00410.

## REFERENCES

1. Kataeva, Z.V. Shilovskikh, O.V., Ulyanov, A.N., Titarenko, E.M.: The results of wavefront aberrations analysis in patients with IOL suture fixation to the iris // IRTC Eye Microsurgery Ekaterinburg Center, Ekaterinburg, Russia. pp.~12 (2019).
2. MI 2083--93. Recommendations: The State System for Providing Unification of Measurements. Definition of Measuring Results and Estimation of their Errors. Moscow, Gosstandart, Russia. Official edition (1993).
3. R 40.2.028--2003. Recommendations: The State System for Providing Unification of Measurements. Recommendations of Building the Calibration Characteristics. Estimation of Errors (Uncertainties) of Linear Calibration Squares Mean Method. Moscow, Gosstandart, Russia. Official edition (2003).
4. Mann, H.B., Whitney, D.R.: On a test of whether one of two random variables is stochastically larger than the other // [Annals of Mathematical Statistics](#). vol. 18. P. 50--60 (1947).
5. Gubler, E.V., Genkin, A.A.: Application of non-parametric statistical criteria in medicine-biological investigations. Medicine, Leningrad (1973) (in Russian).
6. Smirnov, N.V., Dunin-Barkovsky, I.V.: Course in probability theory and mathematical statistics for technical applications. Nauka, Moscow (1969) (in Russian).
7. Cramer, H.: Mathematical Methods of Statistics. Mir, Moscow, (1975) (in Russian).
8. Student: The probable error of a mean. // [Biometrika](#). vol. 6 (1). P. 1-25 (1908).
9. Rockafellar, R., Tyrrell, W., Roger, J-B.: Variational Analysis. Springer-Verlag (2005).